Putting Broca's region into context

fMRI evidence for a role in predictive language processing

Kristensen, Line Burholt; Wallentin, Mikkel

Published in:
Cognitive Neuroscience of Natural Language Use

Publication date:
2015

Document version
Early version, also known as pre-print

Citation for published version (APA):
Kristensen, L. B., & Wallentin, M. (2015). Putting Broca's region into context: fMRI evidence for a role in predictive language processing. In R. M. Willems (Ed.), *Cognitive Neuroscience of Natural Language Use* (pp. 160-181). Cambridge University Press.

Please cite as: Kristensen, Line Burholt & Wallentin, Mikkel (in press) Putting Broca's region into context – fMRI evidence for a role in predictive language processing. In Willems, R. (ed.): *Cognitive Neuroscience of Natural Language Use*, Cambridge University Press.

Putting Broca's region into context – fMRI evidence for a role in predictive language processing.

Line Burholt Kristensen(1,2), Mikkel Wallentin (2,3)

1. Department of Scandinavian Studies and Linguistics, University of Copenhagen, Njalsgade 120, DK-2300 Copenhagen S, Denmark

2. Center of Functionally Integrative Neuroscience, Aarhus University Hospital, Nørrebrogade 44, Building 10G, 5th floor, DK-8000 Aarhus C, Denmark

3. Center for Semiotics, Aarhus University, Building 1467, Jens Chr. Skous Vej 7, DK-8000 Aarhus C, Denmark

Please cite as: Kristensen, Line Burholt & Wallentin, Mikkel (in press) Putting Broca's region into context – fMRI evidence for a role in predictive language processing. In Willems, R. (ed.): *Cognitive Neuroscience of Natural Language Use*, Cambridge University Press.

Abstract

Broca's region is known to play a key role in speech production as well as in the processing of language input. Still, the exact function (or functions) of Broca's region remains highly disputed. Within the generativist framework it has been argued that part of Broca's region is dedicated to syntactical analysis. Others, however, have related Broca's region activity to more domain-general processes, e.g. working memory load and argument hierarchy demands. We here present results that show how contextual cues completely alter the effects of syntax in behaviour and in Broca's region, and suggest that activation in this area reflects general linguistic processing costs or prediction error. We review the fMRI literature in the light of this theory.

Please cite as: Kristensen, Line Burholt & Wallentin, Mikkel (in press) Putting Broca's region into context – fMRI evidence for a role in predictive language processing. In Willems, R. (ed.): *Cognitive Neuroscience of Natural Language Use*, Cambridge University Press.

Introduction: The controversy over Broca's region

In 1861 Paul Broca presented the brain of one of his patients to the anthropological society in Paris. Before his death, this patient had displayed a severe speech deficit, being unable to say more than a single word, "Tan", while apparently maintaining many of his other mental faculties (Broca, 1861). Broca found that the patient had a large lesion in the brain's left inferior frontal gyrus (LIFG). Since then, this area, now often referred to as Broca's region, has been considered a key speech/language brain region. With the advent of cell staining techniques, Korbinian Brodmann (Brodmann, 1909) found that the LIFG, based on the cytoarchitecture, could be subdivided into distinct regions: Brodmann areas 44, 45 and 47 (BA 44/45/47). The subregions are depicted in figure 1. Agrammatical speech was already early considered to be a specific symptom in aphasiology (e.g. Kussmaul, 1877), and it has subsequently been argued that Broca's region plays a significant role in the processing of syntax, both in comprehension and production of language (Friedmann, 2006; Grodzinsky & Santi, 2008).

The exact definition of Broca's region and its subregions has been the subject of later controversy. Usually, BA44 and BA45 are considered the core regions. Some definitions of the region include neighboring areas such as BA 47 and part of BA 6 (Hagoort, 2005), though receptorarchitectonic studies indicate that BA 47 is distinct from BA 44/45 (see Amunts & Zilles, 2006; 2012 for a discussion). Leaving aside these definitorial concerns, we will here be primarily concerned with the function of BA 44/45.

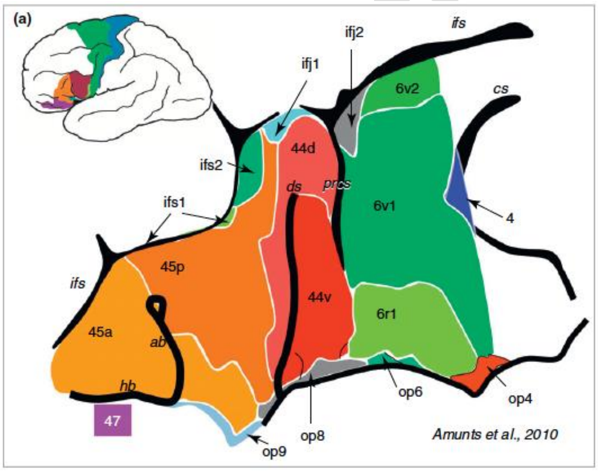

Figure 1. Map of Broca's region

Broca's region map based on the distribution of receptors of neurotransmitters and modulators. BA44 and BA45 were found to be very similar in structure (with some subdivisions), whereas BA47 was found to be very distinct. Reprinted with permission from the authors (Amunts & Zilles, 2012, figure 4)

The association between BA 44/45 and syntactic processing is recognized by most neurolinguistic theories, and the function of Broca's region is rarely defined without reference to sentence processing studies. It is, however, an intensely debated question how the relation between syntactic processing and Broca's region should be modeled (Angela D. Friederici, Rüschemeyer, Hahne, & Fiebach, 2003; Grodzinsky & Santi, 2008; Hagoort, 2005). The question is highly controversial because it stirs up two basic and longstanding oppositions within theoretical linguistics. Linguists both disagree on the definition of *syntax* and *syntactic processing*, and on the relation between language and other cognitive functions.

The generativist transformation-based flank (Ben-Shachar, Palti, & Grodzinsky, 2004; Grodzinsky, 2000; Grodzinsky & Santi, 2008) subscribes to the idea that "[…] *language is a distinct, modularly organized neurological entity*" (Grodzinsky, 2000:1). This idea takes its point of departure in Chomsky's (1965) distinction between different kinds of innate language components, e.g. a phonological component, a semantic component and a syntactic component. The generativist paradigm sees Broca's region as a highly specialized separate linguistic module dealing with the transformational component, a subcomponent of the syntactic component (Grodzinsky, 2000).

Functionalist-cognitivist paradigms within linguistics (e.g.Dik, 1997; Engberg-Pedersen, Fortescue, Harder, Heltoft, & Jakobsen, 1996; Van Valin & LaPolla, 1997), on the other hand, see linguistic structure (including syntactic structures) as shaped by cognition and usage, i.e. by the way we communicate. Functionalist approaches do not see syntactic processes as self-contained, but as dependent on domain-general cognitive processes, e.g. working memory and prediction (Chater & Manning, 2006). Broca's region thereby becomes not the seat of an isolated syntactic module, but of some (possibly domain-general) function that happens to serve a central function in language and communication (Kaan & Swaab, 2002).

In short, the bone of contention is whether syntactic structure is independent or functionally motivated, and whether linguistic processes are based on self-contained modules or imply high-level non-linguistic cognitive functions. When defining the relation between syntactic processing and Broca's region, both of these questions must be dealt with.

The discussions of Broca's region have been fueled with results from a variety of studies: from early lesion studies (Wernicke, 1874) via studies involving symptom/lesion mapping (D. Caplan, Waters, Dede, Michaud, & Reddy, 2007; D Caplan, et al., 2007; Christiansen, Louise Kelly, Shillcock, & Greenfield, 2010; Dronkers, Wilkins, Van Valin, Redfern, & Jaeger, 2004; Friedmann, 2006), behavioral studies (K. j. Christensen, J. Kizach, & A. Nyvad, 2013) and computational models of language processing (Levy, 2008; Spivey & Tanenhaus, 1998) to ERP studies (Bornkessel, Schlesewsky, & Friederici, 2003; A D Friederici, Pfeifer, &

Hahne, 1993; Hagoort, Wassenaar, & Brown, 2003; Hahne & Friederici, 1999) and neuroimaging studies (Ben-Shachar, et al., 2004; Bornkessel, Zysset, Friederici, von Cramon, & Schlesewsky, 2005; K. R. Christensen, J. Kizach, & A. M. Nyvad, 2013; Christensen & Wallentin, 2011; Christian J. Fiebach, Schlesewsky, Lohmann, von Cramon, & Friederici, 2005; C. J. Fiebach & Schubotz, 2006; Grewe, et al., 2005; Hagoort, Hald, Bastiaansen, & Petersson, 2004; Kim, et al., 2009; Makuuchi, Grodzinsky, Amunts, Santi, & Friederici, 2013; Wallentin, Roepstorff, Glover, & Burgess, 2006). However, while the generativist and the functionalist approaches are mutually exclusive in theory, their predictions of single sentence processing, as we shall demonstrate here, often point in the same direction. For a more differentiated description of the function of Broca's region, we therefore suggest including more factors from natural language processing, by examining the role of contextual factors within a predictive coding framework (Chater & Manning, 2006; Clark, 2013; Friston, 2010).

Syntactic processing – different theories with similar predictions

Syntactic processing is often examined by contrasting subject-before-object (SO) structures with object-before-subject (OS) structures. A generativist framework sees the SO structure as basic, e.g. *Mary hit Peter.* The OS structure *Peter, Mary hit* is seen as a transformation of the SO version and therefore as syntactically more complex (Radford, 2004). The generativist association between transformational (sub)processes and Broca's area seems to correspond with empirical findings from a number of language experiments. In accordance with the generativist idea that transformation increases processing demands, OS clauses are, in a number of languages, more difficult to read (Finnish: Hyönä & Hujanen, 1997; Japanese: Miyamoto & Takahashi, 2004) and understand (German: Haupt, Schlesewsky, Roehm, Friederici, & Bornkessel-Schlesewsky, 2008) than SO ones. In accordance with the generativist association between Broca's region and the transformational component, a number of studies have shown more activity in Broca's region for the processing of OS clauses compared to SO ones (Hebrew: Ben-Shachar, et al., 2004; Japanese: Kim, et al., 2009).

Functionalists may also consider OS structures as more complex than SO structures, but they explain the complexity differently. The movement explanation is viewed with skepticism, and the allegation of a transformational component is seen as redundant – accounts that focus on domain-general cognitive processing demands are preferred. A functionalist approach is in line with a number of different explanations to the increased processing demands for OS structures, including the working memory account, the unification model, the argument hierarchy account, and the prediction error account. According to the working memory explanation, OS clauses increase working memory demands as the object cannot be immediately integrated into the sentence structure and must therefore be maintained in

working memory (D. Caplan & Waters, 1999; Christian J. Fiebach, et al., 2005). According to the unification model, the complexity lies in an increased unification load, i.e. the phonological, morphological or syntactical constituents of the sentence are difficult to unify (Hagoort, 2005). The argument hierarchy account sees the semantic relations between the subject and object constituent as the key to the processing difficulties – processing demands are low when the agent appears before the patient (following the alleged order of the argument hierarchy), but in OS clauses the order is usually reversed, increasing argument hierarchy demands (Bornkessel, et al., 2005). Finally, the prediction error account proposes that OS clauses are less predictable to the recipient than SO clauses, e.g. because OS clauses are less frequent. The low frequency of OS clauses therefore should evoke a prediction error due to the difference between the expected input and the input that occurred (Kristensen, Engberg-Pedersen, & Wallentin, in review).
In short, the generativist explanations see the correlation between activity in Broca's area and processing of OS clauses as due to transformational demands, i.e. restricted to linguistic components. The functionalist explanations refer to theories of communicatively grounded or general cognitive demands, e.g. increased working memory demands, increased unification demands, increased argument hierarchy demands or prediction error.

While these explanations are theoretically different, the results that they predict are often hard to tell apart. According to all explanations, SO clauses are easiest to process – they involve no transformations, working memory demands and unification demands are low, the argument hierarchy is rarely violated, and they are more frequent than OS clauses and thereby more predictable. This allows generativists to claim that neural activity in Broca's region correlates with demands on the transformational component, while at the same time functionalists can argue that it correlates with domain-general demands. The problem thus is that the difference between SO and OS clauses cannot be reduced to a difference in word order. SO clauses and OS clauses also differ when it comes to argument hierarchy demands, assumed working memory demands and a number of frequency measures (cf. e.g. Grewe, et al., 2005 for a discussion). While the clustering of sentence-internal factors makes it difficult to manipulate them separately, context manipulations offer a window for testing the hypotheses of transformation-based and usage-based explanations. To examine whether Broca's region is the neural basis of a syntactic language-internal component or of domain-general processes (or of both: Federenko, Duncan, & Kanwisher, 2012), we will argue that it is fruitful to go beyond traditional single-sentence processing and examine sentence processing when it is contextually embedded.

Syntactic processing in context – different theories with different predictions

When examining sentence processing in context, we can make a clearer distinction between the predictions of generativist transformation-based theories and the various kinds of functionalist predictions. According to a transformation-based theory, the activity in Broca's region should be stable across different kinds of context, as transformations are not context-dependent. According to a functionalist approach, linguistic structure is not self-contained, and the increase in activation of Broca's area for OS clauses may well reflect non-linguistic context-dependent factors. When seeing linguistic structure as affected by communicative practices, it makes sense that context has an effect on the processing of word order, as it has on other aspects of linguistic processing (coherence in question-answer pairs: R. Caplan & Dapretto, 2001; coherence between sentences: Kuperberg, Lakshmanan, Caplan, & Holcomb, 2006; coherence between title and paragraph: St George, Kutas, Martinez, & Sereno, 1999). The predictions are therefore different: According to functionalist theories, context may affect activation of Broca's region, and according to generativist theories, it should not.

Two experiments on the effect of discourse context on syntactic processing

The effect of context on syntactic processing is seen in two recent studies of sentence processing in Danish (Kristensen, Engberg-Pedersen, & Poulsen, 2013; Kristensen, et al., in review). Danish is well-suited for examining word order processing, as both objects and subjects can occur in the first position of Danish clauses. Danish is a V2 language (like e.g. German and Swedish), meaning that the finite verb occurs in the second position of the clause, while a variety of constituent types can occur in first position. The following two examples show that both subjects and objects can occur in the first position:

1) Danish SO sentence:

   Hun elsker ham

   she love.PRS him[1]

   'She loves him'

2) Danish OS sentence:

   Ham elsker hun

   him love.PRS she

   'Him, she loves'

Both sentences occur in Danish, though OS sentences are less frequent (Boeg Thomsen & Kristensen, unpublished; Kristensen, 2013; Mikkelsen, unpublished) and limited to certain contexts (Hansen & Heltoft, 2011; Harder & Poulsen, 2001). Example 1) and 2) are disambiguated due to case-marked pronouns. However, not all transitive declarative sentences in Danish contain case-marked pronouns. Some sentences are therefore ambiguous with respect to the distribution of syntactical roles:

3) *Susan elsker Peter*
   Susan love.PRS Peter
   'Susan loves Peter' or 'Susan, Peter loves'

The above ambiguous example can either be interpreted as subject-initial (with *Susan* as the subject) or as object-initial (with *Susan* as the object). While ambiguous clauses do occur in Danish, transitive clauses are typically disambiguated by means of case-marked pronouns (Boeg Thomsen & Kristensen, unpublished; Kristensen, 2013), as in example 1) and 2), by the position of non-finite verbs and (to some extent) of sentential adverbs as well as semantics, e.g. verb argument restrictions or violation of argument-hierarchy (see below for examples). Besides these cues, contextual cues are a further means of disambiguation. While both object and subject can occur in first position in Danish, an object is only licensed in first position if it has a special pragmatic status, e.g. if the object is the most topic-worthy constituent in the sentence (Kristensen, 2013). One way of establishing a referent as topic-worthy is to explicitly contrast it with other elements of a set. In the first of the two sentences below, Ringo is contrasted to the remaining members of The Beatles. In the second sentence, Ringo is thus highly topic-worthy, and the object is licensed in first position.

4) *Susan elsker The Beatles, undtagen Ringo. Ringo hader hun.*
   Susan love.PRS The Beatles, except Ringo. Ringo hate.PRS she
   'Susan loves The Beatles, except Ringo. Ringo, she hates'

The licensing of subjects in first position is less restricted. Subjects can occur in first position even if they are part of an all-focus sentence where all the constituents contain new information, e.g. if the sentence occurs out of context. According to these regulations, SO clauses have a processing advantage over OS clauses when the sentence occurs out of a discourse context.

Context appropriateness affects comprehension

In a reading study, Kristensen et al. (2013) examined whether Danish OS sentences were more easily understood when they occurred in a supportive discourse context, in this case a discourse context that contrasted the referent of the fronted object with other members of a set (as in example 4)). By registering the responses to simple comprehension questions, the study compared the comprehension accuracy rates for unambiguous SO and OS main clauses presented in context. Half of the sentences were shown in a supportive discourse context, the other half was shown in an unsupportive context. The sentences did not contain case-marked pronouns, but were disambiguated by means of the varying positions of non-finite verbs and sentential adverbs. An example of a supportive discourse context for an OS target is given in 5), and the target is given in 6):

5) *Denne historie handler om Anne. Peter brød sig ikke om de andre piger.*
   This story deal.PRS about Anne. Peter liked 3.REFL not about the other girls
   'This story is about Anne. Peter did not like the other girls.'

6) *Anne ville Peter dog invitere til festen.*
   Anne would Peter however invite to party.DEF
   'Anne, however, Peter would invite to the party'

The supportive context in 5) presents both the subject and the object of the target sentence and establishes a set of alternatives to the contrasted element (*Anne* vs. *the other girls*). While the supportive context for *object*-initial target clauses supports a contrastive reading of the *object*, the supportive context for *subject*-initial target clauses (not shown here, but see example 9) supports a contrastive reading of the *subject*.

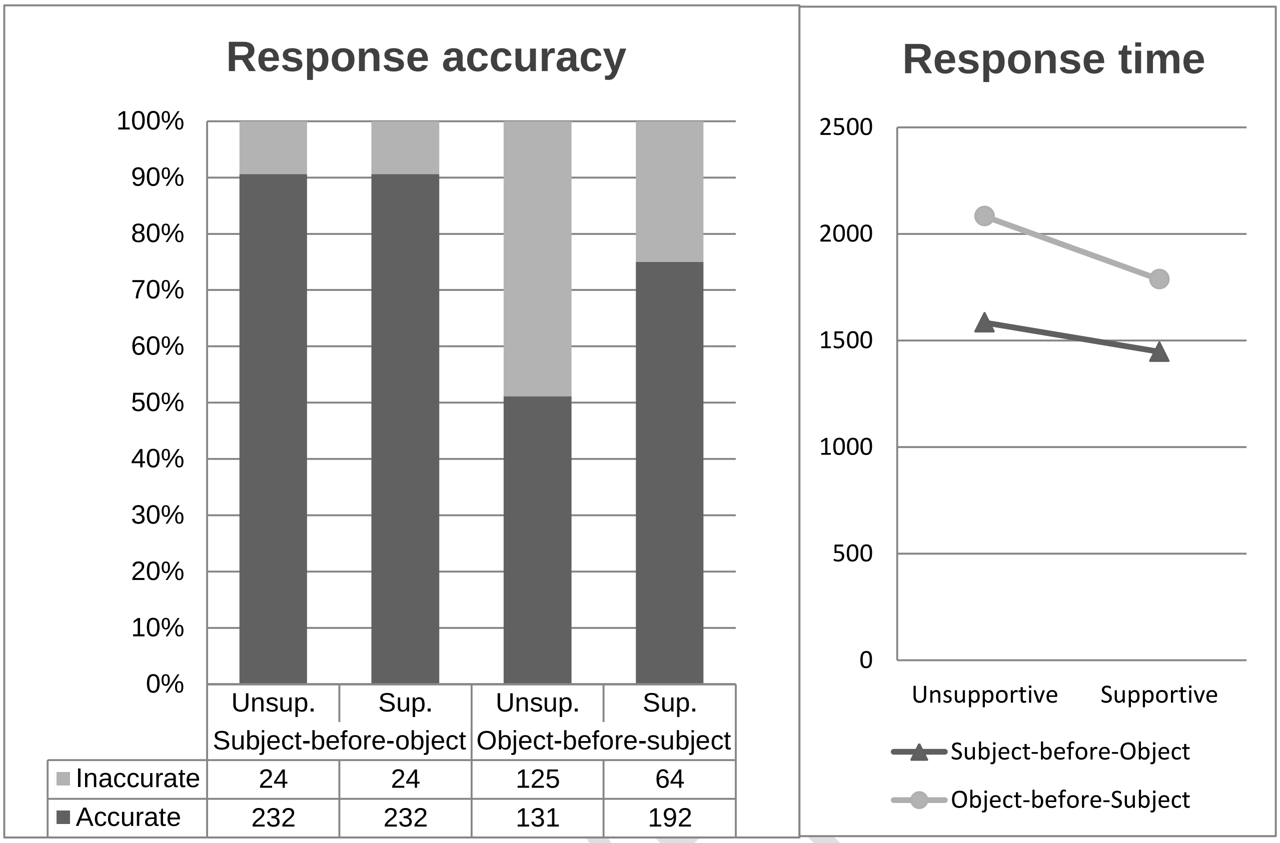

Figure 2. Response accuracy and response time from reading study

In a reading study, Kristensen et al. (2013) found that context supportiveness interacted with word order both for comprehension accuracy and for response time. Object-initial clauses were found to be more context-sensitive than subject-initial ones, as context had a larger facilitating effect on the comprehension of object-initial clauses.

The study examined processing differences between these two kinds of constituent order with or without supportive context. As expected, Kristensen et al. (2013) found greater overall processing difficulties for OS sentences: Responses to comprehension questions were slower for OS sentences than for SO sentences, and responses were more frequently incorrect. However, the study also showed a significant improvement of comprehension when OS sentences occurred in supportive discourse context. In a supportive discourse context, the improvement in response accuracy and response time was more pronounced for OS sentences than for their SO counterparts.

Context appropriateness affects Broca's region

Based on this interaction between context supportiveness and word order for comprehension in the reading time experiment, Kristensen et al. (in review) carried out a neuroimaging experiment in order to investigate whether context supportiveness would affect Broca's region. This study (which will be described in further detail below) found an interaction between context and word order in the activity of BA 44/45 (see figure 3C).

Stimuli for the neuroimaging experiment were auditory and consisted of a control task and a main task. The control task compared the processing of OS and SO clauses *out of context*, i.e. each sentence was presented in isolation. The main task of the neuroimaging experiment studied the processing patterns *in context* using stimuli similar to the behavioural experiment of Kristensen et al. (2013). The main differences between the designs were that stimulus sentences in the neuroimaging experiment were spoken rather than written, that the target sentences involved case-marked pronouns (similar to the sentences in 1) and 2)) instead of proper nouns and that the context manipulation was slightly different. The supportive discourse context still aimed at contrasting the first constituent of the target clause with a set, as seen in the sequence consisting of 7) followed by 8)

7) *Peter overså alle butikstyvene – undtagen Anne.*
Peter overlooked all shoplifters.DEF – except Anne
'Peter overlooked all the shoplifters – except Anne.'

8) *Hende bemærkede han.*
Her noticed he
'Her, he noticed'

The unsupportive context to an OS clause supported the interpretation that the subject of the target, rather than the object, was contrasted with a set of alternatives, e.g. the combination of the context in 9) followed by the target in 8) is pragmatically inappropriate.

9) *Alle overså Anne og hendes bror – undtagen Peter.*
Everybody overlooked Anne and her brother – except Peter
'Everybody overlooked Anne and her brother – except Peter.'

The context in 9) contrasts Peter with a set of people (*alle* = everybody) that overlook Anne and her brother. Anne is part of a group consisting of Anne and her brother, but Peter is singled out as a contrasted constituent. When the context in 9) precedes the target sentence in 8), the subject constituent of the target (*han*, referring to Peter) is the most topic-worthy constituent, and thereby expected to appear in first position. The object (*hende*) is lower-ranking. Compared to the subject (*han*), the object is less likely to be interpreted as involving a contrast or as involving linkage. In this inappropriate context, the object is therefore dispreferred in first position. Similarly, an SO target sentence such as 10) would be preferred after 9), but dispreferred after 7), making the design a balanced factorial design:

10) *Han bemærkede hende.*
he noticed her
'He noticed her'

Focusing on Broca's area in an ROI analysis, the results of the control task (without context) and the main task (with an appropriate vs. inappropriate context) were analysed. The control task showed increased activation in the left BA 45 for OS sentences compared to SO sentences. However, the main task showed an effect of context appropriateness – for appropriate combinations of discourse context and target word order, such as the sequences 7) + 8) and 9) + 10), the BA 44/45 showed decreased activation compared to inappropriate combinations like 7) + 10) and 9) + 8). The study also found an interaction between context supportiveness and word order in BA 44/45: a supportive context had a larger effect on the processing of OS sentences than on SO sentences, i.e. a supportive context led to a larger decrease in activation in BA 44/45 for OS sentences than for SO sentences. This indicates that these parts of the IFG are not restricted to syntactic processing functions, but modulated by several types of language-related expectations (see figure 3A,B,C,D).

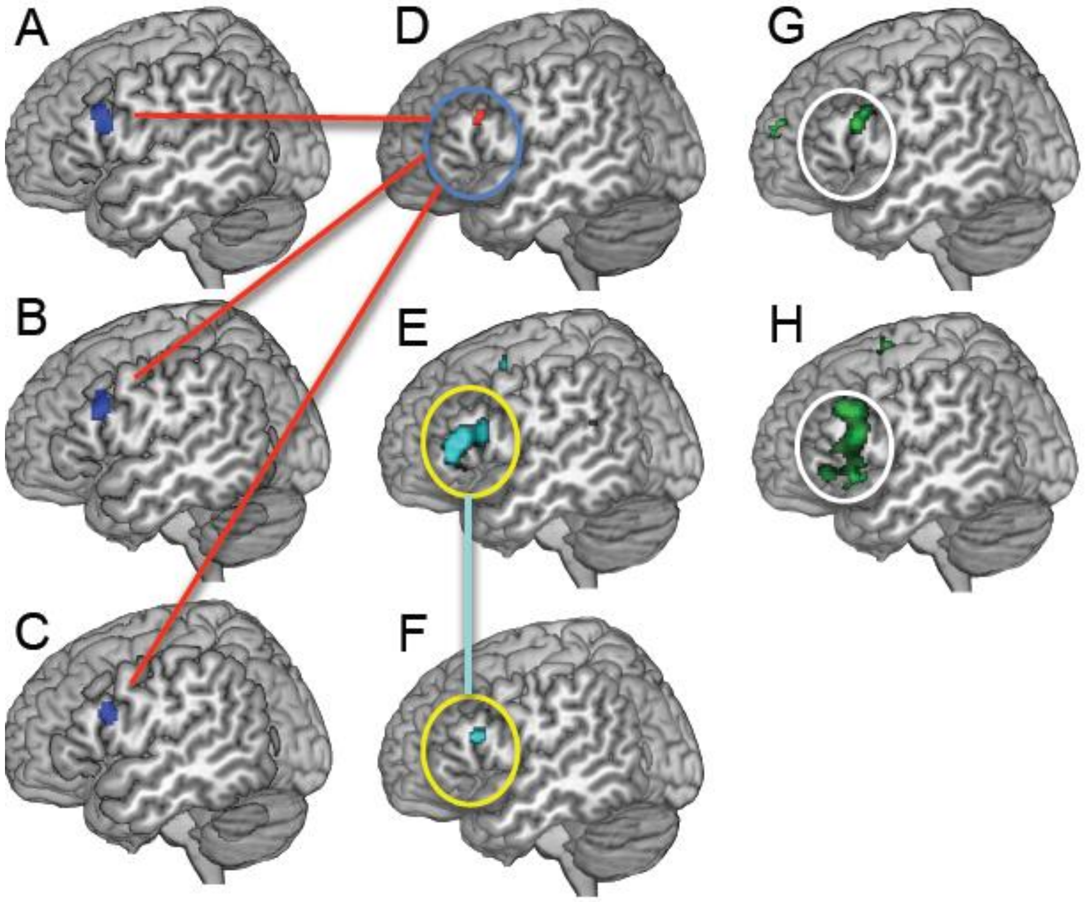

Figure 3. Effects in Broca's area

A: Kristensen et al. (in review) investigated context effects on sentence processing (P<0.05, FWE corrected using a Broca's area ROI). In the control task, a main effect of object-initial word order was found in BA 45 in the absence of contextual cues. B: In the main task, an inappropriate context was found to yield greater response in BA45, as well. C: An interaction between word order and context was also found in the main task, suggesting that object-initial sentences are more context sensitive than subject-initial sentences. D: An overlap for A+B+C was found in BA 45. E+F: Single sentence data from Christensen & Wallentin (2011). Participants made comprehension judgments on single sentences. Average acceptability across participants was used as a covariate in the analysis and it was found that lower acceptability yielded higher activity in BA 45 (E). Response times were also used in the analysis and it was found that higher RT yielded greater activity in BA 45 as well (F). Display contrasts are shown for P<0.0001, uncorrected, t>4.49). G: Wallentin et al. (2006) investigated linguistic referencing of a previously seen image. Their primary focus was the difference between spatial and non-spatial referencing, but they also found that BA45 activity within subjects was positively correlated with response time, regardless of type of reference (figure thresholded at p<0.0001, uncorrected, t>4.7). Broca's area activation went up whenever the participant for some reason had to spend more time on recalling the referenced context. H: Wallentin, Weed et al. (2008) investigated linguistic reference to a previous sentence. Again their primary focus was between different types of reference, but based on F & G we conducted a new analysis with RT as covariate and found an identical effect (data reported here for the first time). Activity in BA 45 and BA47 was positively correlated with trial response time (figure thresholded at P<0.0001, uncorrected, t>4.7, but Broca's area peaks are also significant at FDR-corrected thresholds). These effects were found to be bilateral.

Predictive coding: an alternative functional interpretation

The effects of discourse context supportiveness thus challenge an understanding of Broca's region function as being attributed to transformations alone. Neither can the effects be solely attributed to increased working memory demands or to increased argument hierarchy demands. As intraclausal factors were kept stable in the target sentences of Kristensen et al. (in review), the function of Broca's region in this study cannot be attributed to intraclausal factors such as differences between the subject and the object when it comes to e.g. animacy, given-new status or plausibility of agenthood. If these findings are to be reconciled, the activity observed has to originate in a more abstract function. A part of the functional linguistic paradigm involves the modelling of language processing with probabilistic means (Clark, 2013; Levy, 2008). This predictive coding paradigm entails that the brain extracts statistically stable features from the environment and uses those to predict upcoming stimuli in a hierarchical fashion, i.e. at different levels of abstraction (Friston, 2010). If the predicted input, e.g. a predicted phoneme within a word or a predicted word within a sentence, differs from the actual occurring input, then a prediction error signal is generated. This prediction error causes the predictions to change, both at the current level and further up the predictive hierarchy of abstractions (Chater & Manning, 2006). Ultimately, the predictions will be different the next time the person encounters a similar situation. We suggest interpreting the activity of Broca's region as an indicator of prediction error in the linguistic domain. When the recipient fails to predict the argument order of an upcoming clause, a surprisal effect (Clark, 2013; Levy, 2008) occurs, i.e. the less predictable the sentence, the greater the surprisal, and the more fMRI activation in Broca's region.

When it comes to the word order of upcoming linguistic input, there are a number of possible sources for making predictions about the input, e.g. combinations of speaker characteristics (van Berkum, Brown, Zwiserlood, Koojiman, & Hagoort, 2005), the frequency of the structure both in the ongoing discourse (local frequency, i.e. priming: Pickering & Ferreira, 2008) and in earlier discourse (global frequency), verb restrictions, perceptual cues and the semantics and pragmatics of the context. OS clauses are, as mentioned earlier, relatively infrequent in Danish, and they can therefore be seen as generally less predictable than SO clauses. If Danish language users base their predictions of an upcoming sentence on global frequency (in combination with other previously mentioned sources), they will likely expect sentences to be subject-initial rather than object-initial. However, under specific contextual circumstances, such as those presented above, pragmatics may influence the predictability, and the odds for an OS clause increase drastically. Our findings thus seem to be consistent with an interpretation of Broca's region activity, as indexing a sort of linguistic prediction error. In the following we will review the fMRI literature on Broca's region activation in the light of this hypothesis.

Broca's area and frequency

Frequencies of linguistic input have a strong effect on performance and brain signal. As for global frequency, word frequency effects are found even at the single word level, such as in lexical decision tasks when research participants are asked to classify letter strings as words or as nonwords (Allen, Lien, Smith, Grabbe, & Murphy, 2005; Balling & Baayen, 2012; Forster & Chambers, 1973; Grainger, 1990; Whaley, 1978). Low-frequency words take longer to categorize as words than high-frequency words. In fMRI-studies of lexical decision tasks, low frequency alone was enough to yield an increased Broca's region activation (Christian J Fiebach, Friederici, Müller, & Von Cramon, 2002; Kronbichler, et al., 2004).

While global frequency of structure on its own is unlikely to explain all processing differences between OS and SO clauses (Ferreira, 2003), it is possible to reinterpret a large portion of reading time studies on syntactical manipulations (e.g.Kaiser & Trueswell, 2004) along these lines, i.e. the higher the global frequency of a structure, the shorter the reading time.

Frequencies can also be relevant at shorter time scales. It has been known for a long time that Broca's region activation decreases for word generation tasks if the task is repeated (Raichle, et al., 1994), and even a very rare word order will be less surprising if it is repeated within a short time span. This effect is called structural priming (Bock & Griffin, 2000). In neuroimaging studies, structural priming effects have been found to occur in Broca's region (along with left middle temporal regions), i.e. the activation decreases when a particular linguistic structure is repeated (BA 44/6: Menenti, Gierhan, Segaert, & Hagoort, 2011; BA 44/45: Weber & Indefrey, 2009).

Broca's area and within-sentence contextual effects

Another way of looking at predictability is via contextual bindings that make certain words and readings more expected than others within a particular sentence. For words with multiple meanings (e.g. *bank*), one word meaning may be dominant in relation to the other and thus the predicted reading in absence of a disambiguating context. Zempleni and coworkers (2007) studied this in Dutch with fMRI. When a sentence ended with a word that prompted for the subordinate interpretation of the sentence-initial ambiguous word, Broca's region activity was increased, compared to a sentence that ended with a word prompting for the dominant reading. This is in effect a study of different levels of "cloze probability" (Taylor, 1953) for a given word in a sentence, i.e. the probability with which a reader will continue a given sentence with a given word. In 11) *cow* will have high cloze probability whereas *goat* or *bank account* will have lower cloze probabilities.

11) *The farmer milked his ...*

Similar to the semantic ambiguity resolution study, Obleser and Kotz (2010) found that Broca's region activation was negatively correlated with cloze probability, again suggesting that Broca's region activation relates to linguistic predictability. Similarly, Broca's region activation has also been found to go down as participants adapt to novel metaphors (Cardillo, Watson, Schmidt, Kranjec, & Chatterjee, 2012).

Predictability and acceptability may also be related. Highly predictable sentences are those that we most readily accept as meaningful, whereas less predictable are those where a greater proportion of listeners have difficulties understanding or processing the sentence. Christensen and Wallentin (2011) investigated this in an experiment where participants both read and heard sentences that could be either semantically incongruent or not. They used the so-called locative alternation constructions:

12) *He throws snow on the door.*
13) **He throws the door with snow.*
14) **He blocks rocks on the road.*
15) He blocks the road with rocks.

One construction only works with verbs that focus on the process, e.g. *throw*. The other only works with verbs that focus on the result, e.g. *block*. Participants judged whether a sentence made sense or not. The study design contained a syntactic manipulation as well (sentence 13 and 15 are thought to be more complex than 12 and 14). Both the syntactically more complex sentences and the semantically incongruent sentences yielded a greater Broca's region response with peaks in the exact same region. This suggests that a failure to integrate constituents within a sentence increases Broca's region activity. Further, when the authors looked at the effect of acceptability, they found a second order relationship between the response time of the acceptability question and the number of participants who had rated a particular sentence as comprehensible. Both the strongly incomprehensible (i.e. 13 and 14) and the clearly comprehensible sentences (12 and 15) yielded a fairly short response time, whereas sentences that received mixed responses (e.g. one sentence contained the Danish word *træ* which means both 'wood' and 'tree'. However, only 'wood' made sense within the context), took longest to evaluate, on average. While this is not exactly surprising given the Zempleni study presented above, the authors also found that Broca's region was linearly correlated with response time for the individual sentences, i.e. more ambiguous sentences were accompanied by greater response in Broca's region (See figure 3), again suggesting that lower expectancy leads to greater Broca's region responses.

Predictions can also apply to the number of constituents in a sentence, e.g. if a transitive verb is presented, a direct object will be expected. If this expectation is not met, e.g. due to an intervening prepositional phrase between the verb and the object, then a prediction error will be produced. Fiebach et al. (2005) investigated this type of prediction in a neuroimaging experiment. Participants read the German versions of 16) and 17):

16) *He asks himself who called the doctor after the accident.*
17) *He asks himself who after the accident called the doctor.*

Meaning integration requires that both the verb, the subject and the object have been introduced. In 16), the subject, verb and object occur in succession, while in 17) there is an intervening constituent (*after the accident*) before the verb and the object. The sentence in 17) can thus be seen as putting more strain on expectations and working memory than 16). Indeed, Fiebach and coworkers found that sentences like 17) yielded greater activation of Broca's region than sentences like 16). Interestingly, it made no difference whether the relative pronoun (in English *who*) was subject or object of the embedded clause, i.e. the processing of the German version of 18) did not yield a greater activation than 16), again suggesting that predictability rather than transformations per se cause Broca's region activity to increase.

18) *He asks himself who the doctor called after the accident*

Working memory and language

Fiebach and colleagues (2005) attribute the observed effect to increased working memory demand. However, we will entertain the hypothesis that working memory can in fact also be described within a predictive framework. A few recent attempts at linking the working memory literature (e.g. Baddeley, 1986, 2003; Wallentin, Kristensen, Olsen, & Nielsen, 2011) with the recent probabilistic approaches to cognition (Tenenbaum, Kemp, Griffiths, & Goodman, 2011) have been made (Brady & Tenenbaum, 2013; Orhan & Jacobs, 2013). Until now, this literature has primarily focused on visual working memory. Standard work on working memory assumes a "slot model" under which the individual has a certain number of "slots" into which memory can be stored, e.g. the number of digits recalled in a digit span task. This model, however, cannot account for the great variability in participants' ability to chunk and store different patterns of stimuli (e.g. the numbers 3-3-3 are easier to remember than 3-8-4) and thus expand working memory capacity. A key feature of the predictive models thus is that they take into account perceptual grouping between retained items and some sort of higher order summary of the stimuli that the perceiver tries to maintain. When trying to remember real world scenes, people encode a visual and a semantic gist of what they experience. A sentence might be thought of as a prototypical unit for such a summary. In a predictive framework, this working memory summary can be seen as affecting predictions of incoming input, including linguistic input. So, when a scene is described in a way that is incompatible with the working memory summary, a prediction error is produced and further processing is needed in order to reach a response. We find that this idea is supported by previous neuroimaging studies of linguistic reference to a previously seen image and linguistic reference to a previously read sentence.

Linguistic reference to a previously seen image

Wallentin and coworkers investigated working memory for visual scenes using linguistic references (Wallentin, Roepstorff, & Burgess, 2008; Wallentin, et al., 2006). Participants were shown an image with three referents, a man, a woman and an object. After the image was removed, participants were asked to recall both spatial and non-spatial aspects of the image. The recall questions were presented using simple linguistic cues, e.g. *Was he in front of her?* or *Was he older than her?* The only change across trials was the personal pronouns (*he*/*she*/*it*) used to refer to individual aspects of the image, i.e. the syntax was identical across trials, and the semantics was more or less confined to referential markers. However, similar to Christensen & Wallentin (2011), Broca's region activation was linearly correlated with response time across trials, suggesting that when no easy match is found between a linguistic cue and memory content, additional processing is necessary, and this processing involves Broca's region (see figure 3G). Importantly, the working memory load (number of remembered items) was constant across all trials. This finding is therefore consistent with an interpretation that a mismatch between maintained working memory content and incoming linguistic cue causes a prediction error and hence increased Broca's region activation.

Linguistic reference to a previously read sentence

In a follow-up study, Wallentin and co-workers investigated linguistic reference to a previously encountered sentence, i.e. again focusing on reference across the sentence boundary (Wallentin, Weed, Østergaard, Mouridsen, & Roepstorff, 2008). The participants read sentences about a man and a woman, and their relative spatial and nonspatial relations (e.g. the Danish version of *With their backs to each other stand an elderly man and a young red haired girl*). Subsequently the participants were probed for these internal relations with questions like *Was he facing her?* or *Was he older than her?* The authors replicated their findings of a distinct dorsal network for spatial references, but does the experiment also yield insights with respect to Broca's region? A reanalysis of the neuroimaging data (not reported in Wallentin, Weed, et al., 2008) using response time as a covariate shows a statistically significant effect in Broca's region (with a peak in BA45, see figure 3H). The effect, reported here for the first time, indicates that whenever a mismatch occurs between working memory content and a question, either due to a degradation of the working memory or due to the question not matching the content, a prediction error is generated causing additional processing.

Discussion and future avenues

We have illustrated the importance of including elements of natural language processing in neuroimaging designs, specifically the importance of monitoring predictability and of including a discourse context when examining syntactic processing patterns. Kristensen et al. (2013) showed that discourse context supportiveness affected comprehension of Danish object-initial clauses, while Kristensen et al. (in review) similarly showed that it altered word order effects in BA 44/45 of Broca's region. Based on these results, we argue that the role of Broca's region can be reinterpreted within a predictive coding framework, i.e. activity increases when there is a discrepancy between the predicted input and the input that occurred. The results of other sentence processing studies point in the same direction: When the recipient experiences sentence processing difficulties (as indicated by increased question response times), the activity in BA 44/45 increases (Christensen & Wallentin, 2011; Wallentin, Roepstorff, et al., 2008; Wallentin, et al., 2006).

As neuroimaging studies have a poor temporal resolution, the time course of discourse effects in Broca's region is not clear. We suggest that a supportive discourse context facilitates the prediction of upcoming input, and thereby decreases the prediction error. An alternative explanation would be that a supportive discourse context facilitates the *reanalysis* of the target sentence, i.e. the context facilitation does not exert its effect in Broca's region until after initial processing of the sentence. Still, we find that the prediction approach is advantageous for a number of reasons:

- **The prediction approach fits results from time course studies**

In a reading experiment with Finnish main clauses, Kaiser & Trueswell (2004) found an interaction between discourse context and word order – while a suitable discourse context did not eliminate the difference in reading time between SO and OS clauses, a suitable discourse context had a larger facilitating effect on reading times for OS clauses than for SO clauses. A similar interaction effect between discourse context and word order was found for the processing of Dutch OS and SO relative clauses (Mak, Vonk, & Schriefers, 2008). In both reading experiments, the effects of context are more likely to be prediction-based effects than reanalysis affects, as the effects occurred online.

Likewise, context is known to influence language-related ERP effects, such as the N400 effect and the P600 effect (Federmeier, 2007; van Berkum, 2010). Still, the association between the influence of context on reading time, on ERP effects and on Broca's region needs to be further investigated.

- **The prediction approach can unify existing sentence processing theories**

As we have argued by reinterpreting the results of previous sentence processing studies, a predictive coding framework can unify and integrate theories of both working memory demands, argument hierarchy demands, structural priming and unification.

- **The prediction approach is not specific to language**

The predictive coding framework is in line with previous linguistic as well as non-linguistic research on the functioning of the brain (Friston, 2010; Tenenbaum, et al., 2011). The prediction-based approach thus has the advantage of explaining language processing in terms of principles shared with other kinds of processing, e.g. predictions of visual non-linguistic input (Bar, 2004, 2007). Furthermore, as the prediction approach describes linguistic and non-linguistic processes on the basis of the same principles, the approach can take into account the influences of e.g. visual input or emotional valence on linguistic processing (van Berkum, 2010)

We have argued that the function of Broca's area is not restricted to syntactic processing as such. This broader perspective does, however, not entail that the role of Broca's region is interpreted as all-encompassing and as covering predictions of all sorts.

The extent and limitations to this predictive system and Broca's regions role in it remains to be studied. We have tentatively talked about a linguistic predictive system, which may or may not function in relative isolation from other cognitive operations. But if we accept that Broca's region has a role in linguistic prediction error monitoring, is its role then confined to language or does it go beyond that, to communicative situations broadly defined, or does it apply to all unpredicted events? Judging from the priming literature, there does seem to be a limitation to what Broca's region responds to. In a number of priming experiments, the combinations of linguistic primes and targets affected Broca's region, whereas priming effects for non-linguistic primes and targets did not affect Broca's region (priming of environmental sounds: Bergerbest, Ghahremani, & Gabrieli, 2004; priming pictures of nonsence objects: Vuilleumier, Henson, Driver, & Dolan, 2002). The difference in location of non-linguistic priming effects thus suggests that the prediction error effect in Broca's region is linguistically grounded. Whether it goes beyond the two-sentence range discussed hitherto, is a question for future research to explore.

References

Allen, P. A., Lien, M. C., Smith, A. F., Grabbe, J., & Murphy, M. D. (2005). Evidence for an activation locus of the word-frequency effect in lexical decision. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 713-721.

Amunts, K., & Zilles, K. (2006). A Multimodal analysis of Structure and Function in Broca's Region. In Y. Grodzinsky & K. Amunts (Eds.), *Broca's Region* (pp. 17-30). Oxford: Oxford University Press.

Amunts, K., & Zilles, K. (2012). Architecture and organizational principles of Broca's region. *Trends in cognitive sciences, 16*, 418-426.

Baddeley, A. D. (1986). *Working Memory*. Oxford: Oxford University Press.

Baddeley, A. D. (2003). Working memory and language: an overview. *Journal of Communication Disorders, 36*, 189-208.

Balling, L. W., & Baayen, R. H. (2012). Probability and Surprisal in Auditory Comprehension of Morphologically Complex Words. *Cognition, 125*, 80-106.

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience, 5*, 617-629.

Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends in cognitive sciences, 11*, 280-289.

Ben-Shachar, M., Palti, D., & Grodzinsky, Y. (2004). Neural correlates of syntactic movement: converging evidence from two fMRI experiments. *NeuroImage, 21*, 1320-1336.

Bergerbest, D., Ghahremani, D. G., & Gabrieli, J. D. E. (2004). Neural Correlates of Auditory Repetition Priming: Reduced fMRI Activation in the Auditory Cortex. *Journal of Cognitive Neuroscience, 16*, 966-977.

Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General, 129*, 177-192.

Boeg Thomsen, D., & Kristensen, L. B. (unpublished). Context needed – Semantic role assignment in Danish children and adults.

Bornkessel, I., Schlesewsky, M., & Friederici, A. D. (2003). Eliciting thematic reanalysis effects: The role of syntax-independent information during parsing. *Language and Cognitive Processes, 18*, 269-298.

Bornkessel, I., Zysset, S., Friederici, A. D., von Cramon, D. Y., & Schlesewsky, M. (2005). Who did what to whom? The neural basis of argument hierarchies during language comprehension. *NeuroImage, 26*, 221-233.

Brady, T. F., & Tenenbaum, J. B. (2013). A probabilistic model of visual working memory: Incorporating higher order regularities into working memory capacity estimates. *Psychological Review, 120*, 85-109.

Broca, P. (1861). Remarques sur le siége de la faculté du langage articulé; suives d'une observation d'aphemie. *Bulletin de la Société Anatomique de Paris, 6*, 330-357.

Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Leipzig: Barth.

Caplan, D., Waters, G., Dede, G., Michaud, J., & Reddy, A. (2007). A study of syntactic processing in aphasia I: behavioral (psycholinguistic) aspects. *Brain and Language, 101*, 103-150.

Caplan, D., Waters, G., Kennedy, D., Alpert, N., Makris, N., DeDe, G., Michaud, J., & Reddy, A. (2007). A study of syntactic processing in aphasia II: Neurological aspects. *Brain and Language, 101*, 151-177.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22*, 77-94.

Caplan, R., & Dapretto, M. (2001). Making sense during conversation: an fMRI study. *NeuroReport, 12*, 3625-3632.

Cardillo, E. R., Watson, C. E., Schmidt, G. L., Kranjec, A., & Chatterjee, A. (2012). From novel to familiar: tuning the brain for metaphors. *Neuroimage, 59*, 3212-3221.

Chater, N., & Manning, C. (2006). Probabilistic models of language processing and acquisition. *Trends in cognitive sciences, 10*, 335-344.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge/Massachusetts: MIT Press.

Christensen, K. j., Kizach, J., & Nyvad, A. (2013). Escape from the Island: Grammaticality and (Reduced) Acceptability of wh-island Violations in Danish. *Journal of Psycholinguistic Research, 42*, 51-70.

Christensen, K. R., Kizach, J., & Nyvad, A. M. (2013). The processing of syntactic islands - An fMRI study. *Journal of Neurolinguistics, 26*, 239-251.

Christensen, K. R., & Wallentin, M. (2011). The locative alternation: Distinguishing linguistic processing cost from error signals in Broca's region. *NeuroImage, 56*, 1622-1631.

Christiansen, M. H., Louise Kelly, M., Shillcock, R. C., & Greenfield, K. (2010). Impaired artificial grammar learning in agrammatism. *Cognition, 116*, 382-393.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*, 181-204.

Dik, S. (1997). *The theory of functional grammar*. Berlin; New York: Mouton de Gruyter.

Dronkers, N. F., Wilkins, D. P., Van Valin, R. D., Redfern, B. B., & Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition, 92*, 145-177.

Engberg-Pedersen, E., Fortescue, M., Harder, P., Heltoft, L., & Jakobsen, L. F. (1996). Content, Expression and Structure: Studies in Danish Functional Grammar. Amsterdam: John Benjamins.

Federenko, E., Duncan, J., & Kanwisher, N. (2012). Language-Selective and Domain-General Regions Lie Side by Side within Broca's Area. *Current Biology, 22*, 2059-2062.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology, 44*, 491-505.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology, 47*, 164-203.

Fiebach, C. J., Friederici, A. D., Müller, K., & Von Cramon, D. Y. (2002). fMRI evidence for dual routes to the mental lexicon in visual word recognition. *Journal of Cognitive Neuroscience, 14*, 11-23.

Fiebach, C. J., Schlesewsky, M., Lohmann, G., von Cramon, D. Y., & Friederici, A. D. (2005). Revisiting the role of Broca's area in sentence processing: Syntactic integration versus syntactic working memory. *Human Brain Mapping, 24*, 79-91.

Fiebach, C. J., & Schubotz, R. I. (2006). Dynamic anticipatory processing of hierarchical sequential events: a common role for Broca's area and ventral premotor cortex across domains? *Cortex, 42*, 499-502.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12*, 627-635.

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. *Brain research Cognitive brain research, 1*, 183-192.

Friederici, A. D., Rüschemeyer, S.-A., Hahne, A., & Fiebach, C. J. (2003). The Role of Left Inferior Frontal and Superior Temporal Cortex in Sentence Comprehension: Localizing Syntactic and Semantic Processes. *Cerebral Cortex, 13*, 170-177.

Friedmann, N. (2006). Speech Production in Broca's Agrammatic Aphasia: Syntactic Tree Pruning. In Y. Grodzinky & K. Amunts (Eds.), *Broca's Region*. Oxford: Oxford University Press.

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews. Neuroscience, 11*, 127-138.

Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language, 29*, 228-244.

Grewe, T., Bornkessel, I., Zysset, S., Wiese, R., von Cramon, D. Y., & Schlesewsky, M. (2005). The Emergence of the Unmarked: A New Perspective on the Language-Specific Function of Broca's Area. *Human Brain Mapping, 26*, 178-190.

Grodzinsky, Y. (2000). The neurology of syntax: Language use without Broca's area. *Behavioral and Brain Sciences, 23*, 1-21.

Grodzinsky, Y., & Santi, A. (2008). The battle for Broca's region. *Trends in cognitive sciences, 12*, 474-480.

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends in cognitive sciences, 9*, 416-423.
Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of Word Meaning and World Knowledge in Language Comprehension. *Science, 304*, 438-441.
Hagoort, P., Wassenaar, M., & Brown, C. M. (2003). Syntax-related ERP-effects in Dutch. *Brain research Cognitive brain research, 16*, 38-50.
Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis. Early automatic and late controlled processes. *Journal of Cognitive Neuroscience, 11*, 194-205.
Hansen, E., & Heltoft, L. (2011). *Grammatik over det Danske Sprog*: Det Danske Sprog- og Litteraturselskab.
Harder, P., & Poulsen, S. (2001). Editing for speaking: First position, foregrounding and object fronting in Danish and English. In *Ikonicitet og Struktur* (pp. 1-22): Netværk for Funktionel Lingvistik, Engelsk Insititut, Københavns Universitet.
Haupt, F. S., Schlesewsky, M., Roehm, D., Friederici, A. D., & Bornkessel-Schlesewsky, I. (2008). The status of subject-object reanalyses in the language comprehension architecture. *Journal of Memory and Language, 59*, 54-96.
Hyönä, J., & Hujanen, H. (1997). Effect of word order and case marking on sentence processing in Finnish: An eye fixation analysis. *Quarterly Journal of Experimental Psychology, 50A*, 841-858.
Kaiser, E., & Trueswell, J. C. (2004). The role of discourse context in the processing of a flexible word-order language. *Cognition, 94*, 113-147.
Kim, J., Koizumi, M., Ikuta, N., Fukumitsu, Y., Kimura, N., Iwata, K., Watanabe, J., Yokoyama, S., Sato, S., Horie, K., & Kawashima, R. (2009). Scrambling effects on the processing of Japanese sentences: An fMRI study. *Journal of Neurolinguistics, 22*, 151-166.
Kristensen, L. B. (2013). *Context, you need. Experimental approaches to information structure processing.*, University of Copenhagen.
Kristensen, L. B., Engberg-Pedersen, E., & Poulsen, M. (2013). Context improves comprehension of fronted objects. *Journal of Psycholinguistic Research*.
Kristensen, L. B., Engberg-Pedersen, E., & Wallentin, M. (in review). Context influences word order predictions in Broca's region.
Kronbichler, M., Hutzler, F., Wimmer, H., Mair, A., Staffen, W., & Ladurner, G. (2004). The visual word form area and the frequency with which words are encountered: evidence from a parametric fMRI study. *NeuroImage, 21*, 946-953.
Kuperberg, G. R., Lakshmanan, B. M., Caplan, D. N., & Holcomb, P. J. (2006). Making sense of discourse: An fMRI study of causal inferencing across sentences. *NeuroImage, 33*, 343-361.
Kussmaul, A. (1877). Disturbances of speech. In H. von Ziemssen (Ed.), *Cyclopedia of the practice of medicine*. New York: William Wood.
Kaan, E., & Swaab, T. Y. (2002). The brain circuitry of syntactic comprehension. *Trends in cognitive sciences, 6*, 350-356.
Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106*, 1126-1177.
Mak, W., Vonk, W., & Schriefers, H. (2008). Discourse structure and relative clause processing. *Memory & Cognition, 36*, 170-181.
Makuuchi, M., Grodzinsky, Y., Amunts, K., Santi, A., & Friederici, A. D. (2013). Processing Noncanonical Sentences in Broca's Region: Reflections of Movement Distance and Type. *Cerebral Cortex, 23*, 694-702.
Menenti, L., Gierhan, S. M. E., Segaert, K., & Hagoort, P. (2011). Shared Language: Overlap and Segregation of the Neuronal Infrastructure for Speaking and Listening Revealed by Functional MRI. *Psychological Science, 22*, 1173-1182.
Mikkelsen, L. (unpublished). Verb-second structures.
Miyamoto, E., & Takahashi, S. (2004). Filler-gap dependencies in the processing of scrambling in Japanese. *Language and Linguistics, 5*, 53-166.
Obleser, J., & Kotz, S. A. (2010). Expectancy Constraints in Degraded Speech Modulate the Language Comprehension Network. *Cerebral Cortex, 20*, 633-640.

Orhan, A. E., & Jacobs, R. A. (2013). A probabilistic clustering theory of the organization of visual short-term memory. *Psychological Review, 120*, 297-328.

Pickering, M. J., & Ferreira, V. S. (2008). Structural Priming: A Critical Review. *Psychological Bulletin, 134*, 427-459.

Radford, A. (2004). *Minimalist Syntax*: Cambridge University Press.

Raichle, M. E., Fiez, J. A., Videen, T. O., MacLeod, A. M., Pardo, J. V., Fox, P. T., & Petersen, S. E. (1994). Practice-related changes in human brain functional anatomy during nonmotor learning. *Cerebral Cortex, 4*, 8-26.

Spivey, M. J., & Tanenhaus, M. K. (1998). Syntactic Ambiguity Resolution in Discourse: Modeling the Effects of Referential Context and Lexical Frequency. *Journal of Experimental Psychology: Learning, Memory and Cognition, 24*, 1521-1543.

St George, M., Kutas, M., Martinez, A., & Sereno, M. I. (1999). Semantic integration in reading: engagement of the right hemisphere during discourse processing. *Brain, 122*, 1317-1325.

Taylor, W. L. (1953). "Cloze procedure": a new tool for measuring readability. *Journalism Quarterly, 30*, 415-433.

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: statistics, structure, and abstraction. *Science, 331*, 1279-1285.

van Berkum, J. J. A. (2010). The brain is a prediction machine that cares about good and bad - Any implication for neuropragmatics? *Italian Journal of Linguistics, 22*, 181-208.

van Berkum, J. J. A., Brown, C. M., Zwiserlood, P., Koojiman, V., & Hagoort, P. (2005). Anticipating Upcoming Words in Discourse: Evidence from ERPs and Reading Times. *Journal of Experimental Psychology: Learning, Memory and Cognition, 31*, 443-467.

Van Valin, R. D., & LaPolla, R. J. (1997). *Syntax. Structure, meaning and function*. Cambridge: University Press.

Vuilleumier, P., Henson, R. N., Driver, J., & Dolan, R. J. (2002). Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. *Nature Neuroscience, 5*, 491-499.

Wallentin, M., Kristensen, L. B., Olsen, J. H., & Nielsen, A. H. (2011). Eye movement suppression interferes with construction of object-centered spatial reference frames in working memory. *Brain and Cognition, 77*, 432-437.

Wallentin, M., Nielsen, A. H., Vuust, P., Dohn, A., Roepstorff, A., & Lund, T. E. (2011). Amygdala and heart rate variability responses from listening to emotionally intense parts of a story. *NeuroImage, 58*, 963-973.

Wallentin, M., Roepstorff, A., & Burgess, N. (2008). Frontal eye fields involved in shifting frame of reference within working memory for scenes. *Neuropsychologia, 46*, 399-408.

Wallentin, M., Roepstorff, A., Glover, R., & Burgess, N. (2006). Parallel memory systems for talking about location and age in precuneus, caudate and Broca's region. *Neuroimage, 32*, 1850-1864.

Wallentin, M., Weed, E., Østergaard, L., Mouridsen, K., & Roepstorff, A. (2008). Accessing the mental space - Spatial working memory processes for language and vision overlap in precuneus. *Human Brain Mapping, 29*, 524-532.

Weber, K., & Indefrey, P. (2009). Syntactic priming in German-English bilinguals during sentence comprehension. *NeuroImage, 46*, 1164-1172.

Wernicke, C. (1874). *Der aphasische Symptomencomplex*. Breslau: Cohen & Weigart.

Whaley, C. P. (1978). Word—nonword classification time. *Journal of Verbal Learning and Verbal Behavior, 17*, 143-154.

Zempleni, M.-Z., Renken, R., Hoeks, J. C. J., Hoogduin, J. M., & Stowe, L. A. (2007). Semantic ambiguity processing in sentence context: Evidence from event-related fMRI. *NeuroImage, 3*, 1270-1279.

[1] Abbreviations used for glossing of examples: DEF = definite; PRS = present tense; REFL = reflexive